Assessment of a Markov logic model of crop rotations for early crop mapping

Julien Osman, Jordi Inglada, Jean-François Dejoux

Assessment of a Markov logic model of crop rotations for early crop mapping

Email address: jordi.inglada@cesbio.eu (Jordi Inglada)

Assessment of a Markov logic model of crop rotations for early crop mapping

Julien Osman[a], Jordi Inglada[a,*], Jean-François Dejoux[a]

[a] *CESBIO - UMR 5126*
18 avenue Edouard Belin
31401 TOULOUSE CEDEX 9 - France

Abstract

Detailed and timely information on crop area, production and yield is important for the assessment of environmental impacts of agriculture, for the monitoring of the land use and management practices, and for food security early warning systems. A machine learning approach is proposed to model crop rotations which can predict with good accuracy, at the beginning of the agricultural season, the crops most likely to be present in a given field using the crop sequence of the previous 3 to 5 years. The approach is able to learn from data and to integrate expert knowledge represented as first-order logic rules. Its accuracy is assessed using the French Land Parcel Information System implemented in the frame of the EU's Common Agricultural Policy. This assessment is done using different settings in terms of temporal depth and spatial generalization coverage. The obtained results show that the proposed approach is able to predict the crop type of each field, before the beginning of the crop season, with an accuracy as high as 60%, which is better than the results obtained with current approaches based on remote sensing imagery.

Keywords: early crop type mapping, crop rotations, Markov Logic Networks

*Corresponding author
Email address: jordi.inglada@cesbio.eu (Jordi Inglada)

1. Introduction

Detailed and timely information on crop area, production and yield is important for the assessment of environmental impacts of agriculture (Tilman, 1999), for the monitoring of the land use and management practices, and for food security early warning systems (Gebbers and Adamchuk, 2010). Yield production can be forecasted using models which need information about the surface covered by each type of crop (Resop et al., 2012).

There are different ways of gathering this information, such as statistical surveys or automatic mapping using Earth observation remote sensing imagery. Statistical surveys are expensive to implement, since they need field work, which is time consuming when large areas need to be covered. The use of remote sensing imagery has been found to produce good quality maps when using high resolution satellite image time series (Inglada and Garrigues, 2010). These approaches use supervised classification techniques which efficiently exploit satellite image time series acquired during the agricultural season. Describing the approach used for the supervised classification of satellite images is beyond the scope of this paper and the details can be found in (Inglada and Garrigues, 2010), (Petitjean et al., 2012) or (Petitjean et al., 2014).

As an example of these approaches, figure 1 presents a 5-class crop map obtained using a time series of 13 images acquired by the Formosat-2 satellite during 2009 over a study site near Toulouse in Southern France. The data set is described in Osman et al. (2012). The supervised classification is performed using a Support Vector Machine as described in Inglada and Garrigues (2010).

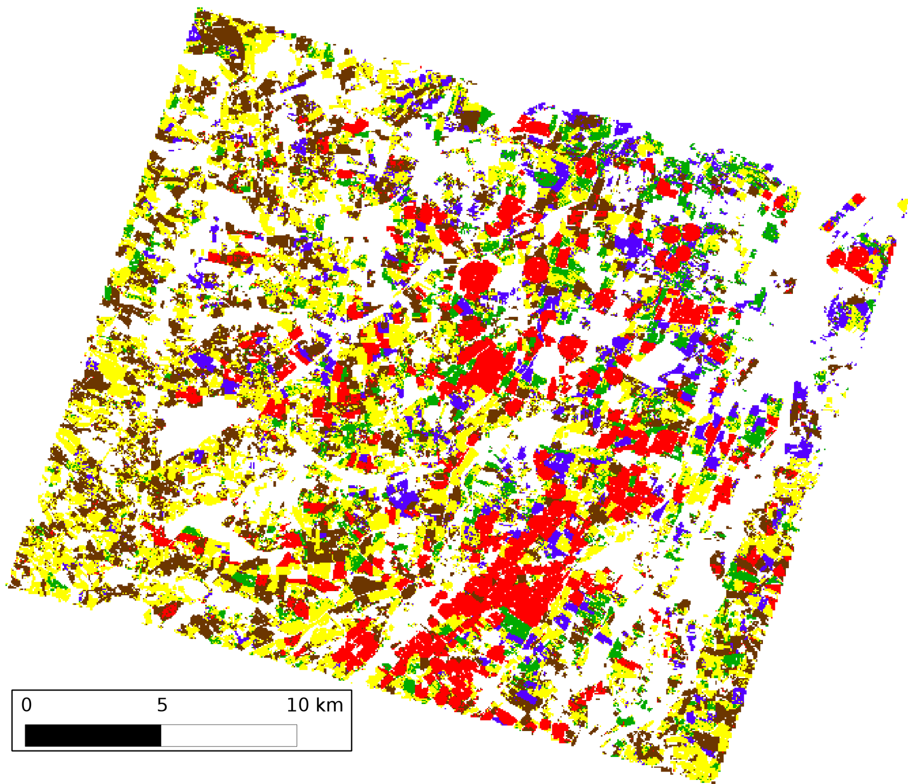


Figure 1: Example of crop map obtained by supervised classification of satellite image time series. Only croplands are classified. Corn (red), wheat (yellow), rapeseed (purple), barley (green), sunflower (brown). White areas represent non croplands.

The resulting classification has an accuracy close to 90%. However, this accuracy can only be achieved at the end of the agricultural season when all images are available. This delay in crop map production has led the remote sensing community to develop near-real-time approaches, where the maps are updated during the season every time a new image is available. Figure 2 shows the evolution of the accuracy of each map produced during the season. A point in the curve represents the accuracy obtained using all the images available up to a given date. In this particular example, one can observe that a quality close

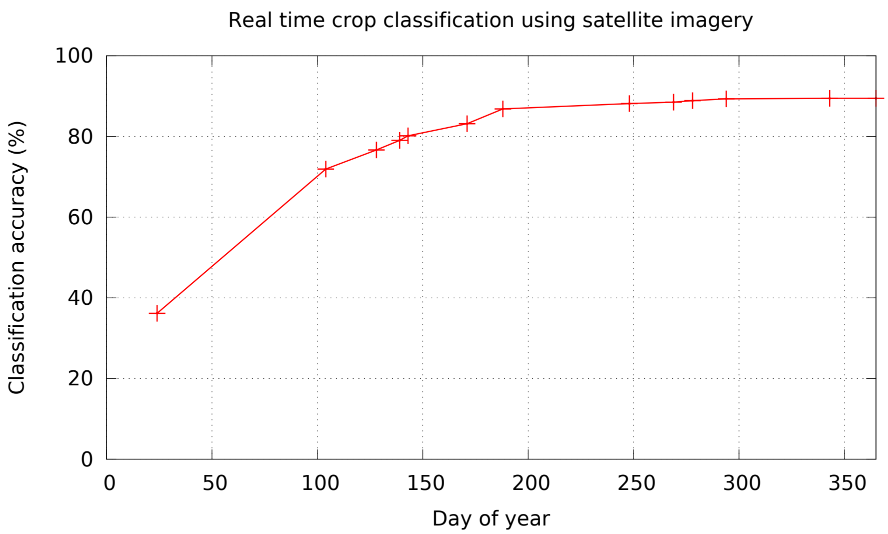


Figure 2: Classification accuracy obtained with satellite image time series. Each cross represents a new image acquisition. The accuracy increases when more images are available.

to the maximum can be obtained before 200 days into the year. However, no information is available before the first image is acquired at the end of January. For many crop systems, the beginning of the season coincides with the end of Autumn or the beginning of Winter. In this period, satellite images are very likely to be cloudy and therefore of little use for crop mapping. Furthermore, the accuracy of the land cover classification obtained with only one image is below 40%, which is not enough for most applications.

The goal of this paper is to introduce an approach which is able to produce land cover maps for agricultural areas at the beginning of the crop season without relying on remote sensing imagery. We propose to use the knowledge about the crop type which was present in every field the previous seasons to predict the crop grown the current year. The proposed approach uses a statistical model for crop rotations.

Crop rotations – specific sequences of crops in successive years – improve or maintain crop yield while reducing input demands for fertilizers and pesticides, and therefore they are widely used by farmers. This regularity on the agricultural practices allows predicting with some accuracy the type of crop present in a given field at one point in time if the previous crop sequence is known.

Many crop rotation models exist, ranging from purely agronomic (crop-soil simulation models (Wechsung et al., 2000)), to approaches integrating expert knowledge and field data (Dogliotti et al., 2003). The complexity of these models makes them difficult to adapt to variable situations and evolving conditions. Crop rotations may evolve in time, either slowly due to for instance climate change impact in rain-fed crops, or very quickly due to environmental regulations dealing with the use of pesticides or water management. Economic factors, as for instance seed prices, can also introduce drastic changes. Hence, crop rotation models which can be easily updated and which can exploit the history of the different territories are needed.

Yearly cropland mapping can be obtained either using farmers administrative declarations or maps produced using remote sensing data at the end of the season (like the one of figure 1). Therefore, the history of the fields can be known.

We propose a machine learning approach to model crop rotations which can predict, at the beginning of a season, with good accuracy, the crops the most likely to be present in a given field, using the crop sequence of the previous 3 to 5 years.

We assess its accuracy using the French Land Parcel Information System RPG in different settings in terms of temporal depth and spatial generalization coverage.

The paper is organized as follows. In section 2, we review several approaches for crop rotation modeling in the literature. Section 3 presents the proposed approach. In section 4, we present the type of data on which our approach relies and we define the experimental setup used for this work; then, we present the details of the assessment and analyze the results. The paper ends with a conclusion and some perspectives.

2. Modeling crop sequences

The predictive model presented in this work (section 3) aims at providing a first guess of crop type maps before the beginning of the crop season. Our model uses knowledge about crop rotations.

Crop rotations have been intensively studied by both agronomists and economists leading to farm management models in the economics and life sciences models in agronomy. Some of them are presented in section 2.1.1. They often require inputs of sequences of crops grown on a specific field over several years. In recent years, there has been an increased focus on sustainable farming systems. This has led to an increase in the use of farm models used to assess the environmental impact of farming. In models of complete exploitations including crop production, it is important to consider the rotation of crops, since this has a major impact on the consequences of the crop production.

However, for the forecasting of crop type mapping, there are specific needs which are not covered by existing modeling approaches. These specific needs are:

1. Field level information: the crop type has to be predicted for every individual field; aggregate data or regional trends are not enough.
2. Different landscapes and different climatic conditions lead to different management practices. Therefore, regional information has to be combined with field-level history.
3. The approach should be portable to different countries and regions of the globe with minimum adaptations. Therefore, it should be able to both, learn from data and to exploit *expert knowledge*. The approach should also be able to use only one of these 2 types of information in case the other one is not available.
4. To cover very large areas, the approach must not rely exclusively on field surveys which are expensive in terms of time and manpower.
5. The model should be able to evolve in time to take into account changing conditions which influence managing practices (climate change, regulatory constraints).

To the best of our knowledge, no existing approach in the literature allows fulfilling all these requirements.

2.1. Existing approaches in the literature

Crop rotation modeling has been addressed in different ways. We may classify these approaches in 2 groups:

1. the approaches using mainly theoretical knowledge, that is models from life sciences, economics or using expert knowledge by agronomists;
2. the approaches which learn from data.

2.1.1. Theoretical knowledge

One simple example of theoretical knowledge is the ROTAT software tool (Dogliotti et al., 2003) which generates all possible rotations of the crops present in a particular area, and then applies a selection based on agronomic criteria provided by experts. This approach allows producing accurate results at the exploitation level, but not at the field level.

The creation of transition matrices adapted to the agricultural landscape under study requires expert knowledge on the type of crop rotation to model and an understanding of the internal dynamics of crop successions. Such knowledge may be derived from research on decision-making by farmers about crop succession (Castellazzi et al., 2008). Castellazzi et al. use Markov chains with transition probabilities set by experts, but their values are limited to 0 and 1.

The specialization of the models to particular sites needs adequate tools. For example Detlefsen et al. (Detlefsen and Jensen, 2007) propose the use of network modeling to find an optimal crop rotation for a given selection of crops on a given piece of land. This model can give advice about the appropriate crop to be grown on a field, but it needs information about the farm (surface, number of fields) and about the costs of farming operations (ploughing, etc.). This kind of information may not be available when mapping very large areas.

Farm management models often produce average crop shares over a number of years, whereas models from the natural sciences often require inputs of sequences of crops grown on a specific field over several years.

For instance, the SWIM model used by Wechsung et al. (Wechsung et al., 2000) can not be applied efficiently over large areas at the individual field level, since it needs very detailed information about specific parameters of the crops. The works of Klöcking et al. (Klöcking et al., 2003) or Salmon-Monviola et al. (Salmon-Monviola et al., 2012) fall in the category of models which perform stochastic simulations for scenarios, but not for accurate mapping at the field level.

In interdisciplinary modeling, this difference can be an obstacle. To bridge this gap, an approach is presented in (Aurbacher and Dabbert, 2011) that allows disaggregating results from farm management models to the level required by many natural science models. This spatial disaggregation consists in deriving a spatial distribution of some information which is only available as a summary for a large area. Aurbacher et al. (Aurbacher and Dabbert, 2011) use Markov chains for the disaggregation at the field level. This approach needs detailed knowledge about the activity at the field and farm levels, as well as other economical information as for instance gross margin. This level of detail is difficult to obtain for large areas and therefore the approach is not suited to mapping.

The integration of many types of knowledge is challenging, and one of the approaches for overcoming this difficulty is to use multi-agent systems, as for instance in the Maelia platform (Taillandier et al., 2011). This approach suffers from the same drawbacks as the previous ones: the need to access detailed knowledge at the farm level.

The main drawback of models based on theoretical knowledge is their inability to easily adapt to changing conditions, since these new conditions have to be accounted for in the models, or adaptive decision rules have to be implemented. However, some attempts have been made to take into account changes. For instance, Supit et al. (Supit et al., 2012) model climate change impacts on potential and rain-fed crop yields on the European continent using the outputs of three General Circulation Models in combination with a weather generator. However, this model is only able to evolve with respect to climate and not with respect to other types of changes.

2.1.2. Automatic learning from data

One way to overcome the problem of adaptation to changing environments or to specific areas, is to integrate field surveys or similar data in the models.

There are models which are used to describe existing data, as for instance CarrotAge (Le Ber et al., 2006), which allows analyzing spatio-temporal data to study the cropping patterns of a territory. The results of CarrotAge are interpreted by agronomists and used in research works linking agricultural land use and water management. The underlying algorithms use Markov models. The main limitation of CarrotAge for our needs is that it does not perform crop prediction at the field level.

Another example is the crop rotation model CropRota (Schönhart et al., 2011), which integrates agronomic criteria and observed land use data to generate typical crop rotations for farms and regions. CropRota does not work at the field level.

Similar to the previous one, ROTOR (Bachinger and Zander, 2007) is a tool for generating and evaluating crop rotations for organic farming systems. It was developed using data from field experiments, farm trials and surveys and expert knowledge. Its originality is the integration of a soil–crop simulation model. As the two previous approaches, ROTOR does not perform predictions at the field level.

As our goal is to map the croplands, we need not only to model the transitions of crops, but also to take into account the geospatial information available.

Usually, the data available for integration in models comes from census and has no continuous spatial distribution. Many approaches for the spatialization of this kind of information exist, as for instance krigging (Flatman and Yfantis, 1984). In the case of crop distribution, You et al. (You et al., 2006) proposed an approach to go from census data to raster information, but their work is not applied to the field level, which is needed in our case for crop mapping.

Although limited to 3 crops, Xiao et al. (Xiao et al., 2014) used field level information to perform a regional scale analysis, but they did not perform forecasting of the selected crops in the individual fields.

Among the cited approaches, none of them fulfill the 5 constrains listed at the beginning of this section. However, some of these works have shown that statistical modeling of crop rotations in general, and Markovian models in particular are appropriate tools for crop type prediction. The drawback of the Markovian approaches used in the literature is that they are not easily updated when expert knowledge complementary to existing data is available.

3. Modeling with Markov Logic

We start (section 3.1) by justifying the use of Markovian approaches for crop rotation modeling and we point out their main limitation for our needs : the impossibility of easily integrating expert knowledge. We then present in section 3.2 the Markov Logic approach which solves this issue. Finally, in section 3.3 we describe how to use Markov Logic Networks to model crop rotations and to forecast future crops.

3.1. Properties of the model

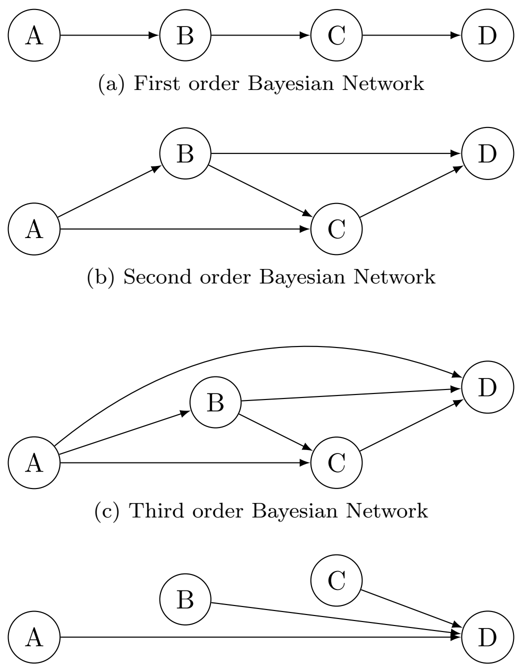
At the beginning of section 2, the specific needs for the forecasting of crops at the field level were listed. After the literature review on crop rotation models, the properties that a model for our application should possess can now be precised.

1. Learning from past sequences, both at the field and at the regional scale. This allows taking into account regional trends together with specific field information.
2. Exploiting the past information for every particular field (either using Land Parcel Information Systems or existing land-cover maps).
3. Incorporating changes in practices without needing the compilation of new data bases containing examples of these evolutions. This allows the model to quickly evolve without the need of a time lag before being able to exploit information about changing conditions.

As we saw in section 2, existing approaches to assess agricultural practices focus on the assessment of single crops or statistical data because spatially explicit information on practically applied crop rotations was lacking (Lorenz et al., 2013), but this is not the case anymore in the EU. For instance Leteinturier et al. (Leteinturier et al., 2006) used the land parcel management system implemented in the frame of EU's Common Agricultural Policy to assess many common rotation types from an agro-environmental perspective. Also, in the USA, the USDA's Cropland Data Layer provides annual crop cover data at 30 m. resolution (Boryan et al., 2011).

When learning from data representing sequential states of variables, the Markovian properties are often used. In a Markovian process, the next state depends only on the current state and not on the sequence of events that preceded it. This allows to efficiently learn the probability of any particular sequence of states by computing only the probability of transition between individual states. As a matter of fact, most of the approaches similar to those presented in section 2.1.2 use these approaches.

One of the most frequently used Markovian models are Bayesian Networks (BN) (Friedman and Koller, 2003; Heckerman et al., 1995) which are today one of the most promising approaches to Data Mining and Knowledge Discovery in databases. A BN is a graph (structure of the network) where each node is a random variable (for instance the crop grown on a particular field on a given year) and each edge represents the degree of dependence between the random variables (the probability of transition between states). Figure 3 illustrates some examples of BN.



(a) First order Bayesian Network

(b) Second order Bayesian Network

(c) Third order Bayesian Network

(d) Common effect Bayesian Network

Figure 3: Examples of Bayesian Networks.

A Markov Random Field, MRF, (or Markovian Network, MN) is similar to a BN in its representation of dependencies (Kindermann et al., 1980); the differences being that BN are directed and acyclic, whereas MN are undirected and may be cyclic. Thus, a MN can represent certain dependencies that a BN cannot (such as cyclic dependencies); on the other hand, it cannot represent certain dependencies that a BN can (such as induced dependencies).

BN and MRF need probability estimates which can be learnt from data. However, they cannot easily incorporate other types of knowledge as for instance logic rules. For instance, in the case of crop rotations, a new regulation about nitrates can change the patterns of the sequences. Changes in prices or a reorientation towards bio-fuel production can lead to yet bigger changes. These expected changes can be expressed with rules, but no data is available for learning until several agricultural seasons have passed. Furthermore, in some cases, the knowledge is easier to express in terms of a set of sentences or formulas in first-order logic (*if-then* rules), rather than in terms of transition probabilities between states. Therefore, an alternative or an extension to BN and MRF is needed.

3.2. Markov Logic

To combine knowledge from databases and knowledge from experts, inference approaches which are able to combine probabilistic learning and rule-based logic reasoning are needed. Combining probability and first-order logic in a single representation has long been a goal of Artificial Intelligence. Probabilistic graphical models like BN make it possible to efficiently handle uncertainty. First-order logic allows to compactly represent a wide variety of knowledge. The combination of probabilistic and propositional models has been one research area of important activity since the mid 1990's (Cussens, 2001; Puech and Muggleton, 2003).

Recently, Markov Logic (ML) (Richardson and Domingos, 2006) was introduced as a simple approach to combining first-order logic and probabilistic graphical models in a single representation. A Markov Logic Network (MLN) is a first-order knowledge base (KB) with a weight attached to each formula[1]. Together with a set of constants representing objects in the domain, it specifies

[1] Logic formulas are also called rules or clauses.

a ground MN[2] containing one feature for each possible grounding of a first-order formula in the KB, with the corresponding weight. Inference in MLNs is performed by Monte Carlo Markov Chains (MCMC) over the minimal subset of the ground network required for answering the query. Weights are efficiently learned from relational databases by iteratively optimizing a pseudo-likelihood measure. Optionally, additional clauses are learned using inductive logic programming techniques. Also, clauses can be added if some prior or expert knowledge is available.

A first-order logic KB can be seen as a set of hard constraints on the set of possible worlds: if a world does not respect one single formula, it has zero probability. In MLN, these constraints are softened: if a world does not verify one formula in the KB it has a lower probability, but not zero. The more formulas a world respects, the more probable it is. Each formula has an associated weight that reflects how strong a constraint is: the higher the weight, the greater the difference in probability between a world that satisfies the formula and one that does not. The weights are not limited in range as probability values are.

Models like MRF and BN can still be represented compactly by MLNs, by defining formulas for the corresponding factors.

Efficient algorithms for learning the structure of the networks and the weights associated to the rules exist (Singla and Domingos, 2005) and they are made available by the authors as a free and open source software implementation (Kok

[2]A ground MN is a MN without free variables in the logic formulas. It is also usually referred as a *possible world*.

et al., 2006) which makes possible the assessment of the approach for our needs.

3.3. The proposed approach

We propose to model each rotation of interest as one rule and use a MLN for the inference. Therefore, the rules do not need to be learned, but only their weights. Using data for a set of years, the weights of each rule are learned. The approach is validated by applying the inference.

The crops of interest for our experiments are wheat, barley, corn, rapeseed and sunflower, which represent 78% of the surface in the study area. The rules are expressed as follows in the case of a 4 year rotation cycle:

$$\{C^a_{n-3}, C^b_{n-2}, C^c_{n-1}\} \rightarrow C^d_n, \omega,$$

which means that the rule which says that a sequence of crop a, followed by crop b, followed by crop c leads to crop d the following year has a weight ω. The notation can be simplified as

$$\{a, b, c, d, \omega\}.$$

The weights ω have to be learned for each possible sequence of crops that has to be modeled. This type of rules corresponds to the same kind of dependency which can be modeled by a common effect BN (figure 3d).

4. Experiments and results

4.1. Description of the available data and the area of study

4.1.1. The French RPG LPIS

The information about the crop rotation used for the assessment of the model was obtained from the *Registre Parcellaire Graphique*, RPG, a topographical

Land Parcel Information System (LPIS) containing the agricultural parcels and the corresponding crops grown.

At the national French level, it contains about 7 million parcels. The system was implemented in 2002 in application of EU directives. It is annually updated by farmers themselves. The information of interest associated to each parcel is:

- the geographical outline of the parcel and an identifier;
- the district where the parcel is located;
- the type of the crop grown a particular year using a 28 class nomenclature;
- the administrative type of the exploitation;
- the age class of the owner for individual owners.

One particularity of the RPG is that the parcels may correspond either to individual fields or to groups of small fields. These groups may be composed by fields where different crops are present. In these cases, the spatial distribution is not given and only the proportion of each crop surface is known.

For the experiments presented here, only individual fields where a single crop is grown were used. This made the analysis easier and the amount of data remained sufficient for the statistical approach to be robust. However, a statistical bias might appear because of the use of a subset of the fields. To solve this issue, techniques have been proposed for the estimation of the spatial distribution of the crops within a group of fields (Inglada et al., 2012) and they could be used in the future.

It is also worth noting that the RPG was used here to have access to a very large geographical area during several years and assess the properties of the proposed model, but other sources of data, as for instance land-cover maps from previous years as the one illustrated in figure 1, could be used without loss of generality.

4.1.2. Study area and time frame

For our study, we used 7 years of data (2006-2012) over a large region in the South of France (figure 4). This amount of data allowed us to assess the model in terms of temporal stability, temporal depth of the rotations as well as spatial homogeneity of the areas.

We used 3 areas of study which are depicted on figure 4:

1. a small area of 20 km$\times$ 20 km (red rectangle) which has rather homogeneous pedo-climatic conditions with about 1700 parcels studied;
2. a medium sized region (dark gray area including the small area) with about 15500 parcels studied and where soils have different types and a sensible North-South climatic gradient is present;
3. a large sized area (light gray area plus the 2 previous ones) with about 72000 parcels studied and presenting a wide variety of soils, landscapes and climatic conditions.

4.2. Experimental setup

4.2.1. Assessment

To assess the capabilities of MLN to give useful information for forecasting the grown crops at the field level, we used the data base presented in section

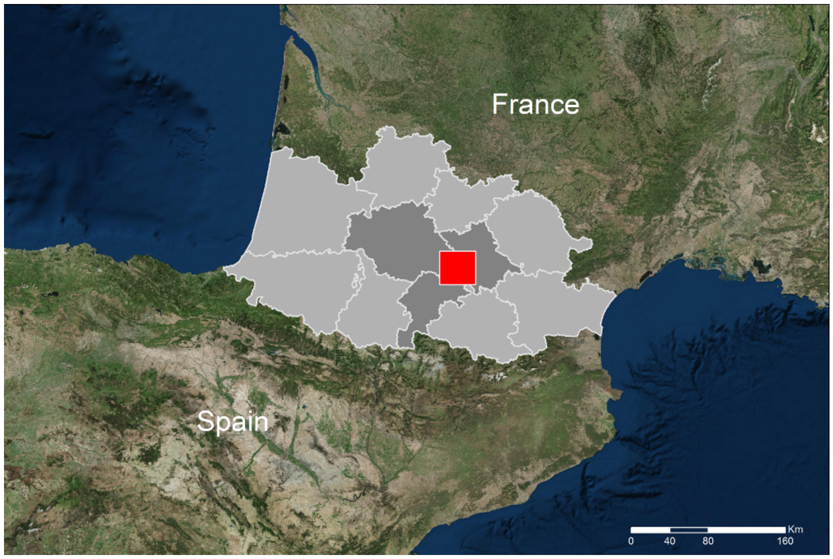


Figure 4: The 3 study regions: in red a 20 km× 20 km area (small), in dark gray the medium area and in light gray the large area.

4.1. We studied the influence of the length of the considered rotations as well as the extent of the area over which the modeling was performed.

To assess the influence of the rotation length, we analyzed 3 different cases: 4 year rotations (that is knowledge of the previous 3 years to forecast the forth one), 5 year rotations and 6 year rotations.

Finally, to assess the impact of the extent of the area (eco-climatic conditions, pedology, etc.), we used the 3 regions presented in figure 4.

4.2.2. Evaluation

To evaluate the quality of the crop prediction, classical tools from the machine learning field were used: the confusion matrix and the Kappa coefficient.

The confusion matrix (also known as contingency table) is a double entry table where row entries are the actual classes (crop in the reference data) and column entries are the predicted classes. Each cell of the table contains the number of elements of the row class predicted by the classifier as belonging to the column class.

The diagonal elements in the matrix represent the number of correctly predicted individuals of each class, i.e. the number of ground truth (reference) individuals with a certain class label that actually obtained the same class label during prediction.

The off-diagonal elements represent misclassified individuals or the classification errors, i.e. the number of ground truth individuals that ended up in another class during classification.

Part of the agreement between the classifier's output and the reference data can be due to chance. The Kappa coefficient (κ) expresses a relative difference between the observed agreement P_o and the random agreement which can be expected if the classifier was random, P_e.

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

where

$$P_o = \frac{1}{n}\sum_{i=1}^{r} n_{ii}$$

is the agreement and

$$P_e = \frac{1}{n^2}\sum_{i=1}^{r} n_{i.}n_{.i}$$

κ is a real number between -1 and 1 and can be interpreted as follows:

Agreement	κ
Excellent	> 0.81
Good	0.80 - 0.61
Moderate	0.60 - 0.41
Weak	0.40 - 0.21
Bad	0.20 - 0.0
Very bad	< 0

4.3. Assessment of the proposed approach

4.3.1. Examples of obtained rotations

To give the reader a sense of the difference between crop rotation frequency and the knowledge modeled by the MLN, the 20 most frequent rotations in the small study area for a 4 year cycle are presented in table 1, and the 20 rules with the highest weights for the same area and the same period are presented in table 2.

In terms of frequency of the rotations, the first thing we note is that the first and the second rotations are the same with a shift of one year. It is interesting to note that these 2 rotations have very high weights in table 2 and these weight are not very different if we take into account that there is a 1.6 ratio in terms of frequency. We can also see that the corn mono-culture is very frequent and the corresponding rule has also a very high weight.

Looking at the first 3 rows of both tables, one may deduce that rule weights yield similar information to frequency of occurrence of rotations. However, this

	2009	2010	2011	2012	number
1	sunflower	wheat	sunflower	wheat	405
2	wheat	sunflower	wheat	sunflower	253
3	corn	corn	corn	corn	113
4	sunflower	wheat	sunflower	barley	46
5	wheat	rapeseed	wheat	rapeseed	46
6	wheat	rapeseed	wheat	sunflower	46
7	wheat	sunflower	wheat	rapeseed	38
8	rapeseed	barley	wheat	rapeseed	34
9	rapeseed	wheat	sunflower	wheat	34
10	sunflower	wheat	rapeseed	wheat	26
11	rapeseed	wheat	rapeseed	wheat	26
12	barley	wheat	sunflower	wheat	26
13	barley	wheat	rapeseed	barley	24
14	sunflower	barley	wheat	sunflower	24
15	wheat	sunflower	barley	sunflower	22
16	sunflower	wheat	wheat	sunflower	21
17	wheat	sunflower	barley	wheat	21
18	barley	wheat	sunflower	barley	19
19	wheat	rapeseed	barley	wheat	18
20	barley	sunflower	wheat	sunflower	16

Table 1: Most frequent rotations in the small area with their corresponding number of occurrences.

is not the case, since the rules represent a conditional probability[3] of the last crop of the sequence with respect to the sequence of the 3 crops which precede it. For instance, rules where corn is present appear in the table (limited to the 20 rules with the highest weights) even if corn is only present in one of the most frequent sequences.

4.3.2. Overview of the behavior

With the data set used, there were 27 different combinations in terms of area, rotation length and particular sets of years. Tables 3, 4 and 5 give an

[3]Although weights are not restricted to the $[0-1]$ intervals as probabilities are. In the same way, the sum of all weights does not have to be 1 as with probabilities. This latter property allows introducing new knowledge not represented in the data when available.

	2009	2010	2011	2012	weight
1	sunflower	wheat	sunflower	wheat	0.752
2	corn	corn	corn	corn	0.699
3	wheat	sunflower	wheat	sunflower	0.601
4	wheat	barley	wheat	barley	0.355
5	corn	corn	rapeseed	barley	0.333
6	corn	corn	rapeseed	rapeseed	0.331
7	corn	rapeseed	corn	rapeseed	0.322
8	corn	corn	rapeseed	wheat	0.319
9	corn	rapeseed	corn	barley	0.317
10	corn	corn	rapeseed	sunflower	0.312
11	sunflower	wheat	barley	sunflower	0.309
12	rapeseed	corn	corn	rapeseed	0.305
13	corn	corn	barley	barley	0.305
14	wheat	barley	wheat	corn	0.304
15	rapeseed	corn	corn	barley	0.302
16	barley	wheat	sunflower	rapeseed	0.302
17	corn	rapeseed	corn	sunflower	0.3
18	corn	barley	corn	barley	0.3
19	wheat	barley	wheat	rapeseed	0.298
20	barley	wheat	sunflower	corn	0.297

Table 2: Higher weight rules in the small area with their corresponding weights ($\{a, b, c, d, \omega\}$).

Small region				
-	2009	2010	2011	2012
4 years	0.51	0.58	0.54	0.60
5 years	-	0.57	0.53	0.61
6 years	-	-	0.54	0.55

Table 3: κ coefficient values for the small region

overview of the results, in terms of κ coefficients, for the small, the medium and the large regions respectively.

The first observation we can make is that most of the κ values were in the high fifties, which is a moderate to good prediction of the crops. It is not surprising to note that the predictions for the small area were the best and those for the large area were the worse, since the eco-pedo-climatic conditions which govern agricultural practices are more homogeneous in the small area. However, the results of the medium area were very close to those of the small area.

In terms of rotation length, we can observe that 4 and 5 years were equivalent for the small and medium regions and that 6 years was worse than 5 which could be explained by the high number of rotations to model in the longer case (4096 combinations with respect to 1024).

Finally, we can observe that the predictions for the year 2011 were the ones with the lower quality independently of the area and of the length of the rotations. This may be explained by the fact that 2009 suffered from an anomalous weather which forced many farmers in the South of France to change the planned winter wheat for a Summer crop like sorghum or sunflower. This modification of practices impacted the statistical representativity of the data.

Medium region				
-	2009	2010	2011	2012
4 years	0.53	0.57	0.51	0.58
5 years	-	0.57	0.52	0.59
6 years	-	-	0.51	0.54

Table 4: κ coefficient values for the medium region

Large region				
-	2009	2010	2011	2012
4 years	0.50	0.56	0.52	0.58
5 years	-	0.50	0.46	0.53
6 years	-	-	0.43	0.43

Table 5: κ coefficient values for the large region

In the following paragraphs, the details of the confusion matrices are analyzed to gain some insight on the behavior of the model.

4.3.3. Area

We focused our interest on the differences of prediction quality between the different regions of different size. In order not to multiply the combinations, we used the results for the length of 5 years and analyzed the confusion matrices which resulted from the averaging the results of the predictions for 3 years (2010 to 2012).

The confusion matrices for the small, the medium and the large areas are presented on tables 6, 7 and 8 respectively.

The first thing we can highlight is that there were no major differences between the small and the medium regions as it was already noted in the overall κ coefficient tables above. The confusion matrices allowed us to check that this stability was reproduced even at the level of the individual crops and their specific confusions.

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	73	5	8	5	9
Corn	5	80	5	4	6
Barley	24	6	32	8	30
Rapeseed	7	5	12	29	46
Sunflower	15	17	28	21	20

Table 6: Confusion matrix for the small region

In terms of confusions, we can see that sunflower was the most difficult crop to predict and more so when the area was very large. In this latter case, the prediction accuracy was lower than random (which would be of 20%). During the past decade, sunflower yields have been steadily decreasing in this region and it is increasingly becoming an opportunity crop to use when the planned winter crop could not be sowed.

At the opposite, wheat and corn were very well predicted and this was mostly because they are the principal crops grown in the area. Rapeseed was much confused with sunflower, since they are usually chosen for economic reasons rather than for agronomic ones. We also see that barley was often predicted as wheat, which is easy to explain because these 2 crops are both straw cereals (and therefore interchangeable form the agronomic point of view) and as stated before, wheat is the most prominent one of those 2. The confusion was stable between areas, but barley was less well predicted when the area was larger mainly because of increasing confusions with rapeseed. The good prediction of corn remained stable independently of the size of the area.

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	74	6	7	5	9
Corn	5	80	4	5	6
Barley	23	9	27	11	30
Rapeseed	7	7	11	26	49
Sunflower	18	18	24	22	18

Table 7: Confusion matrix for the medium region

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	65	8	8	8	10
Corn	6	79	4	5	6
Barley	23	13	22	13	29
Rapeseed	11	10	13	21	45
Sunflower	19	20	24	22	16

Table 8: Confusion matrix for the large region

4.3.4. Length

We limited the study to the medium area and we analyzed the influence of the length of the sequences used for the model (column 2012 of table 4). The results are presented in tables 9, 10 and 11 for the rotations using 4, 5 and 6 years respectively.

The trends that we observe are the following:

- the longest sequences were the most difficult to predict, which is not surprising, since the number of possible combinations is higher and therefore the probability of each one is lower;
- the prediction of corn was good and stable for the different rotation lengths, since most of the corn in the area is grown as mono-culture;
- the prediction of wheat was good but decreased with the length of the sequence;

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	83	4	3	3	7
Corn	5	75	5	4	11
Barley	27	5	12	13	43
Rapeseed	6	7	8	15	64
Sunflower	17	16	17	23	27

Table 9: Confusion matrix for a 4 year sequence.

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	80	4	4	4	7
Corn	5	76	4	7	8
Barley	26	7	28	8	31
Rapeseed	6	8	13	24	49
Sunflower	16	15	25	23	21

Table 10: Confusion matrix for a 5 year sequence.

- rapeseed and sunflower were often confused and their respective prediction accuracies had inverse trends: rapeseed benefited from longer sequences, while sunflower was best predicted with shorter sequences;

- in the previous paragraphs, we observed an important amount of barley being predicted as wheat, and we saw that this confusion diminished when the areas were larger; here we see that this confusion was stable with respect to the length of the sequence, however the prediction of barley benefited from medium length sequences, mainly because the reduction of the confusion with rapeseed.

4.3.5. Simulating drastic changes

In the previous experiments we showed the ability of MLN to predict the crops knowing the past history of the fields. However, from the application point of view, this kind of use is similar to the use of BN, the main advantage

-	Wheat	Corn	Barley	Rapeseed	Sunflower
Wheat	68	8	7	7	10
Corn	7	75	4	6	8
Barley	24	7	25	10	34
Rapeseed	7	7	12	33	41
Sunflower	19	19	15	27	20

Table 11: Confusion matrix for a 6 year sequence.

of MLN being the possibility to have straightforward access to human readable rules instead of having a graphical model which is difficult to interpret when there are many nodes.

However, the use of MLN was proposed because they are able to combine statistical learning with first-order logic rules. This particular property of MLN is interesting to introduce knowledge for which no historical data is available. In the case of early crop mapping, this situation may happen due to new regulations or economic reasons, like seed prices.

Unfortunately, this kind of behavior was not present in our data set, and therefore, we chose to simulate it. The following experiment was carried out. We assumed that for an arbitrary reason, one type of rotation which had been frequent in the past became nearly non existent from a given point in time. We introduced this expected behavior by strongly modifying the weight of the rule related to this particular rotation. We then analyzed how the probability of the crops to be predicted spread among the possible types of crops.

Of course, this kind of event is extreme and not likely to occur as such, but it allowed illustrating the flexibility of the proposed approach.

For this experiment, we used the MLN obtained by performing the training

-	Original	Modified
Corn	0.60	0.0014
Wheat	0.11	0.28
Sunflower	0.11	0.28
Rapeseed	0.088	0.22
Barley	0.089	0.23

Table 12: Predicted probabilities for each crop for the rotation $\{C_{n-3}^{corn}, C_{n-2}^{corn}, C_{n-1}^{corn}\} \rightarrow C_n^d$ with the original weight and the modified one.

on the medium sized region and using the years from 2008 to 2011 (used to predict the crops in 2012).

We chose the sequence $\{corn, corn, corn, corn\}$ whose weight was 0.699 and modified it to have a weight of $-\infty$. It is interesting to note that only this rule was modified. We then analyzed the predicted probability by the MLN for different rotations in the case where we kept the original weight for the rule or we used the modified weight.

Table 12 shows the predicted probability for class d on year n for the rules $\{C_{n-3}^{corn}, C_{n-2}^{corn}, C_{n-1}^{corn}\} \rightarrow C_n^d$ for the original (learned from data) weight and the modified one. As one can see, the original setting predicted *corn* with a probability of 0.6, the other classes having a very low probability. In the case where $\{corn, corn, corn, corn\}$ was nearly non existent, *corn* was predicted with a probability which was practically zero, while the other classes were predicted with similar probability, but those which previously had higher probabilities (wheat and sunflower) still had higher chances than rapeseed and barley.

It is worth noting that no re-learning from the data had to be done, so this kind of changes can be introduced in the model at no cost.

It was also necessary to check that the modification of a particular rule did not have effect on other rules. To verify the correct behavior of the model, we applied the same kind of analysis to other rules. In the case of one of the most frequent rotations of the study area $\{sunflower, wheat, sunflower, wheat\}$, which is described by the rules $\{C_{n-3}^{sunflower}, C_{n-2}^{wheat}, C_{n-1}^{sunflower}\} \rightarrow C_n^d$, there was no modification of the probabilities after changing the weight of the rule $\{corn, corn, corn, corn\}$.

The same behavior occurred for the set of rules $\{C_{n-3}^{wheat}, C_{n-2}^{barley}, C_{n-1}^{wheat}\} \rightarrow C_n^d$. Finally, a family of rules containing 2 consecutive years of corn was not modified either.

In the case of a BN, this modification would have required to modify the training data and learn the transition probabilities again, since it is impossible to modify the probability of a particular sequence of events without modifying all the rest.

The point here is not that the probabilities of the other crops did not change. In a realistic setting, the relative proportion of other crops may evolve due to economic or agronomic reasons. If knowledge about these evolutions is available (for instance, a Summer crop will be replaced by another Summer crop), it can be easily introduced in the model. The main advantage of MLN with respect to other statistical models like BN is that the changes are limited to the particular set of rules directly related to the events and these changes are not propagated to unrelated rules in the model.

5. Conclusions

In this paper we presented a model which allows predicting the crop grown on a field when the crops grown the previous 3 to 5 years are known. This kind of prediction is useful for the production of crop maps at the field level at the beginning of the agricultural season.

Our model applies machine learning techniques using a Land Parcel Information System, or any other kind of land cover maps from previous years, to model crop rotation patterns. With respect to other models existing in the literature, our approach allows combining automatic learning from data with expert knowledge and make predictions at the field level. We have demonstrated with an illustrative example that this property allows introducing constraints that cannot appear in historical data, like for instance new regulations which may change agricultural practices.

We assessed the behavior of the model in terms of scale (area covered) and crop rotation length. We concluded that, in terms of statistical accuracy, the results are good and can be used as a first guess for early crop mapping. The obtained results showed that the proposed approach is able to predict the crop type of each field, before the beginning of the crop season, with an accuracy which can go up to 60%, which is better than the results obtained with current approaches based on remote sensing imagery.

One application of this model would be to use it to complement other techniques for crop mapping as for instance remote sensing image classification. Remote sensing image time series can achieve good results if enough images are available, usually towards the end of the season. The prediction of the most probable crop could allow achieving good results earlier in the season.

The results presented here open perspectives in terms of exploitation of the approach, as for instance including other information as digital elevation models, climatic data or soil type maps.

Acknowledgment

The first author acknowledges the funding by CNES, the French Space Agency, and Région Midi-Pyrénées through a 3 year PhD grant.

Aurbacher, J. and S. Dabbert
2011. Generating crop sequences in land-use models using maximum entropy and Markov chains. *Agricultural Systems*, 104(6):470–479.

Bachinger, J. and P. Zander
2007. ROTOR, a tool for generating and evaluating crop rotations for organic farming systems. *European Journal of Agronomy*, 26(2):130–143.

Boryan, C., Z. Yang, R. Mueller, and M. Craig
2011. Monitoring us agriculture: the us department of agriculture, national agricultural statistics service, cropland data layer program. *Geocarto International*, 26(5):341–358.

Castellazzi, M., G. Wood, P. J. Burgess, J. Morris, K. Conrad, and J. Perry
2008. A systematic representation of crop rotations. *Agricultural Systems*, 97(1):26–33.

Cussens, J.

2001. Integrating probabilistic and logical reasoning. In *Foundations of Bayesianism*, Pp. 241–260. Springer.

Detlefsen, N. K. and A. L. Jensen

2007. Modelling optimal crop sequences using network flows. *Agricultural Systems*, 94(2):566–572.

Dogliotti, S., W. Rossing, and M. Van Ittersum

2003. ROTAT, a tool for systematically generating crop rotations. *European Journal of Agronomy*, 19(2):239–250.

Flatman, G. T. and A. A. Yfantis

1984. Geostatistical strategy for soil sampling: The survey and the census. *Environmental monitoring and assessment*, 4(4):335–349.

Friedman, N. and D. Koller

2003. Being Bayesian About Network Structure. A Bayesian Approach to Structure Discovery in Bayesian Networks. *Machine Learning*, 50(1-2):95–125.

Gebbers, R. and V. I. Adamchuk

2010. Precision agriculture and food security. *Science*, 327(5967):828–831.

Heckerman, D., D. Geiger, and D. Chickering

1995. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning*, 20(3):197–243.

Inglada, J., J. Dejoux, O. Hagolle, and G. Dedieu
2012. Multi-temporal remote sensing image segmentation of croplands constrained by a topographical database. In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, Pp. 6781–6784. IEEE.

Inglada, J. and S. Garrigues
2010. Land-cover maps from partially cloudy multi-temporal image series: optimal temporal sampling and cloud removal. In *IEEE Intenational Geoscience and Remote Sensing Symposium*, Honolulu, Hawaii, USA.

Kindermann, R., J. L. Snell, et al.
1980. *Markov random fields and their applications*, volume 1. American Mathematical Society Providence, RI.

Klöcking, B., B. Ströbl, S. Knoblauch, U. Maier, B. Pfützner, and A. Gericke
2003. Development and allocation of land-use scenarios in agriculture for hydrological impact studies. *Physics and Chemistry of the Earth*, 28(Recent Development in River Basin Research and Management):1311 – 1321.

Kok, S., M. Sumner, M. Richardson, P. Singla, H. Poon, and P. Domingos
2006. The Alchemy system for statistical relational AI (Technical Report). Department of Computer Science and Engineering, University of Washington, Seattle, WA.

Le Ber, F., M. Benoît, C. Schott, J.-F. Mari, and C. Mignolet
2006. Studying crop sequences with CarrotAge, a HMM-based data mining software. *Ecological Modelling*, 191(1):170–185.

Leteinturier, B., J. Herman, F. d. Longueville, L. Quintin, and R. Oger
2006. Adaptation of a crop sequence indicator based on a land parcel management system. *Agriculture, ecosystems & environment*, 112(4):324–334.

Lorenz, M., C. Fuerst, and E. Thiel
2013. A methodological approach for deriving regional crop rotations as basis for the assessment of the impact of agricultural strategies using soil erosion as example. *Journal of environmental management*, 127:S37–S47.

Osman, J., J. Inglada, J. Dejoux, O. Hagolle, and G. Dedieu
2012. Fusion of multi-temporal high resolution optical image series and crop rotation information for land-cover map production. In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International*, Pp. 6785–6788. IEEE.

Petitjean, F., J. Inglada, and P. Gancarski
2012. Satellite image time series analysis under time warping. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(8):3081–3095.

Petitjean, F., J. Inglada, and P. Gancarski
2014. Assessing the quality of temporal high-resolution classifications with low-resolution satellite image time series. *International Journal of Remote Sensing*, 35(7):2693–2712.

Puech, A. and S. Muggleton
2003. A comparison of stochastic logic programs and Bayesian logic programs.

In *Proceedings of the IJCAI-2003 Workshop on Learning Statistical Models from Relational Data*, Pp. 121–129.

Resop, J. P., D. H. Fleisher, Q. Wang, D. J. Timlin, and V. R. Reddy 2012. Combining explanatory crop models with geospatial data for regional analyses of crop yield using field-scale modeling units. *Computers and Electronics in Agriculture*, 89:51–61.

Richardson, M. and P. Domingos 2006. Markov logic networks. *Machine learning*, 62(1-2):107–136.

Salmon-Monviola, J., P. Durand, F. Ferchaud, F. Oehler, and L. Sorel 2012. Modelling spatial dynamics of cropping systems to assess agricultural practices at the catchment scale. *Computers and Electronics in Agriculture*, 81:1–13.

Schönhart, M., E. Schmid, and U. A. Schneider 2011. CropRota–A crop rotation model to support integrated land use assessments. *European Journal of Agronomy*, 34(4):263–277.

Singla, P. and P. Domingos 2005. Discriminative training of Markov logic networks. In *AAAI*, volume 5, Pp. 868–873.

Supit, I., C. Van Diepen, A. De Wit, J. Wolf, P. Kabat, B. Baruth, and F. Ludwig 2012. Assessing climate change effects on European crop yields using the crop growth monitoring system and a weather generator. *Agricultural and Forest Meteorology*, 164:96–111.

Taillandier, P., O. Therond, et al.
2011. Use of the Belief Theory to formalize Agent Decision Making Processes: Application to cropping Plan Decision Making. In *European Simulation and Modelling Conference*, Pp. 138–142.

Tilman, D.
1999. Global environmental impacts of agricultural expansion: the need for sustainable and efficient practices. *Proceedings of the National Academy of Sciences*, 96(11):5995–6000.

Wechsung, F., V. Krysanova, M. Flechsig, and S. Schaphoff
2000. May land use change reduce the water deficiency problem caused by reduced brown coal mining in the state of brandenburg? (english). In *Landscape and urban planning*, volume 51, Pp. 177 – 189.

Xiao, Y., C. Mignolet, J.-F. Mari, and M. Benoît
2014. Modeling the spatial distribution of crop sequences at a large regional scale using land-cover survey data: A case from France. *Computers and Electronics in Agriculture*, 102:51–63.

You, L., S. Wood, and U. Wood-Sichra
2006. Generating global crop distribution maps: from census to grid. In *Selected paper at IAEA 2006 Conference at Brisbane, Australia*, volume 202, Pp. 1–16.